MODEL

Environmental and Resource Economics **14:** 365–383, 1999.

Does Question Format Matter? Valuing an Endangered Species

DIXIE WATTS REAVES[1], RANDALL A. KRAMER[2] and THOMAS P. HOLMES[3]

[1]*Department of Agricultural and Applied Economics, Virginia Tech, Blacksburg, VA 24061-0401, USA (e-mail: dixie@vt.edu);* [2]*Nicholas School of the Environment, Duke University, USA;* [3]*Southern Research Station, USDA Forest Service, USA*

Accepted 14 October 1998

Abstract. A three-way treatment design is used to compare contingent valuation response formats. Respondents are asked to value an endangered species (the red-cockaded woodpecker) and the restoration of its habitat following a natural disaster. For three question formats (open-ended, payment card, and double-bounded dichotomous choice), differences in survey response rates, item non-response rates, and protest bids are examined. Bootstrap techniques are used to compare means across formats and to explore differences in willingness to pay (WTP) distribution functions. Convergent validity is found in a comparison of mean WTP values, although some differences are apparent in the cumulative distribution functions. Differences across formats are also identified in item non-response rates and proportion of protest bids. Overall, the payment card format exhibits desirable properties relative to the other two formats.

Key words: contingent valuation, endangered species, question format, red-cockaded woodpecker

JEL classification: Q26

With a growing interest in ecosystem protection and restoration, economic research has continued to seek improved methods for valuing environmental amenities. One non-market valuation tool, the contingent valuation method (CVM), relies on surveys to ask individuals directly about the value of a specified change in environmental quality. The Department of Energy and the Environmental Protection Agency recently sponsored a workshop with the primary goal of identifying key research issues relating to contingent valuation (CV). One issue that continues to raise concerns for CV practitioners is the choice of question format (elicitation method).

This paper addresses this unresolved issue in contingent valuation using data from a survey on restoration of an endangered species and its habitat. Three question formats (open-ended, payment card, and double-bounded dichotomous choice) are compared using a split-sample design to address several questions: (1) are there differences in response rates, item non-response rates, and proportion of protest bids across question formats? (2) do measures of central tendency suggest conver-

gent validity across the formats? and (3) do the welfare values obtained from the different question formats come from the same underlying distribution?

The paper is organized as follows. Section one contains a discussion of the question formats compared in this study. Section two presents the comparison tests conducted for the three formats. Section three discusses the survey design. Results are presented in section four, and section five concludes and draws implications for future work.

1. Question Formats

The three primary question formats for asking contingent valuation questions are open-ended (OE), dichotomous choice[1] (DC), and payment card (PC). In the OE format of questioning, a respondent is simply asked how much he or she would be willing to pay for an improvement in some environmental good. This format has been criticized as unlikely to provide the most reliable valuations, since respondents are being asked to place a dollar value on a particular public good, and they are not accustomed to such behavior (Federal Register 1993).

In the DC format, respondents are presented with a dollar amount ($\$A$) and asked if they would be willing to pay that amount for the stated environmental improvement. Dollar amounts from a chosen range are randomly assigned to respondents, and respondents state whether or not they would pay the offered amount. A double-bounded dichotomous choice question can be utilized: if the respondent answers yes (no) to the first amount, a second higher (lower) amount is offered. It has been argued that adding a single follow-up question to the original DC question improves the efficiency of the WTP estimate (Hanemann et al. 1991).[2]

In some DC applications, the willingness to pay question is posed as a referendum-type question. Rather than being asked whether or not they would be willing to pay $\$A$ (usually by paying higher prices or by paying into some form of private fund), respondents may be asked if they would vote for a referendum that supports the environmental change, given that it will cost the individual A additional dollars. Whether specifically asked as a referendum question, or merely presented as a dollar value to which respondents answer yes or no, this format is argued to more closely resemble the decisions that respondents make in their daily lives, unlike the valuation decision required by the OE format (Mitchell and Carson 1989).

The dichotomous choice question format has been endorsed by the NOAA Panel (Federal Register 1993) who argued that there is no strategic reason for an individual to answer untruthfully.[3] Disagreeing, Fisher (1994) argues that "the case for closed-ended CV responses being free of strategic bias has not been made either in theory or by empirical findings" (1994, p. 8). Furthermore, Fisher feels that the choice between open-ended and closed-ended format "is currently a very lively issue in the CV literature" (1994, pp. 7–8). In a recent study comparing open-ended and dichotomous choice values, Loomis et al. (1997) suggest that

"it may be premature to abandon use of open-ended WTP questions" (1997, p. 121). Furthermore, they indicate that additional research is needed "before one may determine whether dichotomous choice question formats should always be preferred to open-ended questions in CVM as is recommended by the NOAA panel" (1997, p. 121).

The DC format can be statistically inefficient in that large sample sizes are required to identify the underlying distribution of WTP values. It has also been found that DC questions can induce starting point bias[4] (Mitchell and Carson 1989; McFadden 1994; Cameron and Quiggin 1994; Holmes and Kramer 1995; Herriges and Shogren 1996) and "yea-saying" (Holmes and Kramer 1995; Kanninen 1995; Boyle et al. 1996; Ready et al. 1996). Thus, despite the NOAA panel's endorsement of the DC format, the issue of question format remains unresolved.

The PC format is somewhat of a hybrid of the first two, in that respondents are presented with a range of willingness to pay values, and are asked to circle the one that represents their willingness to pay for the environmental improvement. While not faced with a simple yes or no decision as to whether or not to pay a specified "price", the payment card format does reduce some of the difficulty of trying to assign a dollar value with no guidance. Payment cards were initially developed to avoid the starting point bias problem associated with iterative bidding games but have been associated with problems involving implied value cues (Mitchell and Carson 1986) and range bias (Mitchell and Carson 1989). However, a recent study found the PC format to be free of range and centering biases when the range of the WTP distribution included on the payment card is sufficiently large that it does not constrain the respondent (i.e., the range is not truncated) (Rowe et al. 1996). This result suggests that the range of values displayed on the card should be based on pre-test surveys.

Given that these three question formats have been prevalent in the literature, and given the lack of consensus on the appropriate format to use, this research focuses on a three-way comparison of these formats.

2. Comparison Tests

A number of studies have compared various pairs of the three question formats to determine if they yield similar results (for PC vs. DC comparisons, see Boyle and Bishop 1988; Cameron and Huppert 1991; and Holmes and Kramer 1995; for OE vs. DC comparisons, see Loomis et al. 1997; Johnson et al., 1990; Boyle et al. 1996; Loomis 1990; Kealy and Turner 1993; McFadden and Leonard 1993; and Loomis et al. 1993; and for OE vs. PC comparisons, see Smith and Desvousges 1986). This study provides a unique three-way comparison of question formats, using independent subsamples drawn from the same population.

While most studies have focused on a test of equivalence of the willingness to pay amounts, this study looks at additional comparisons. The willingness to pay (WTP) value is certainly important,[5] but comparisons of other key factors, such as

response rate, item non-response and proportion of protest responses are additional issues to consider when comparing the performance of question formats.

Two possible outcomes could result from comparisons across question formats. First, there could be no significant differences across formats, in which case the question becomes which format is the easiest and most cost effective to use?[6] Alternatively, there could be a significant difference across two or more of the formats. Then one must ask which format, if any, is preferred? By looking at comparisons other than mean willingness to pay, this study attempts to address these questions.

In conducting the three-way comparisons in this study, the following issues are addressed. First, do response rate parameters differ across question formats? If, as has been argued, the dichotomous choice question is easier to answer since it more closely resembles market transactions or true referenda (see, for example, Schumann 1994), then one might expect the response rate for the DC format to be higher than for either the OE or PC format.[7] If a survey is returned, but the willingness to pay question itself is unanswered, then the response is not usable. Therefore, another comparison of interest is that of item non-response. Again, since the dichotomous choice format is alleged to be easier for respondents, item non-response of DC respondents might be expected to be lower than for the other two formats. Finally, if the WTP question is answered with a protest response (such as a protest zero or a protest "no"), the means by which protest responses are treated may have a significant bearing on value estimates. To compare performance in terms of overall survey response, item non-response, and protest responses, z-tests are conducted to test for equal proportions across formats.

The second major issue considered is the distribution of WTP values across question formats. The first moment (mean) of the WTP distribution has been emphasized by researchers concerned with the reliability and validity of contingent valuation measures (Desvousges et al. 1993; Kealy et al. 1990; McFadden and Leonard 1993). As in these studies, paired t-tests of equivalent means are conducted in this study.[8] Standard errors for the OE and PC samples are estimated from the raw data and standard errors for the double-bounded dichotomous choice are estimated using a bootstrap technique.[9]

It is also informative to evaluate whether or not WTP values elicited by the various question formats come from the same underlying distribution. Since both the OE and PC data are continuous[10] by nature, where individuals can state any value along a continuum of values (as opposed to responding yes or no to a discrete value as in the DC format), it is possible to conduct a non-parametric test to determine if the proportions of individuals willing to pay each dollar amount are equivalent for the two formats. One method for testing equivalence is the test of homogeneity of proportions. The null hypothesis, that for each bid level the proportion of respondents stating that amount is equal for the OE and PC question formats, is tested using a chi-squared test statistic that compares observed and expected outcomes.

A second test of the hypothesis that question format has no influence on the distribution of WTP values is conducted by constructing synthetic responses based on the raw data from the three samples. Monte Carlo methods have been used previously to construct synthetic responses to DC questions based on PC (Cameron and Huppert 1991; Holmes and Kramer 1995) and OE (Desvousges et al. 1993) data. To improve the efficiency of the statistical tests and to conduct the tests with a single estimator, data from the three sub-samples (OE, PC, and DC) are pooled using Monte Carlo methods to create synthetic double-bounded DC observations from the OE and PC data. In addition, the sampling variability of the actual double-bounded DC responses is simulated by modifying the procedures described by Duffield and Patterson (1991) to account for the binomial probability distribution of DC payment responses.

Dummy variables are created for the synthetic OE and PC subsamples in the pooled data regression. For each bootstrap replication of the synthetic data, parameters for the double-bounded DC model are estimated and stored. Standard errors of the vector of bootstrap parameter estimates are computed and used to test the null hypothesis that parameter estimates on the OE and PC dummy variables in the pooled regression are not different from zero (i.e. the question format does not influence the distribution of WTP values).

3. Survey Design

Analyses are performed on data from a valuation study of the red-cockaded woodpecker (RCW), listed as endangered in 1970. The RCW is an endangered species that is not largely known to the general population and has non-consumptive use value to a very small segment of the population. Consequently, nonuse values comprise most of the total value.

An important population of the bird was severely damaged when Hurricane Hugo struck the Francis Marion National Forest (FMNF) in 1989. The 250,000-acre FMNF, established in 1936, contained 75,000 acres of old-growth longleaf pine, the preferred habitat of the RCW. The RCW population in the FMNF was the densest, the second largest, and the only known increasing population in the world. As a result of the hurricane, most of the woodpeckers' cavity trees for nesting and roosting were destroyed (87%), much of its foraging habitat was damaged (50–60%), and many of the birds were killed (63%) (Hooper, Watson and Escano 1990).

A CVM survey was undertaken by the authors to estimate the value of protection and restoration activities for the bird and its habitat. The unfortunate natural disaster provided a unique opportunity to measure individuals' willingness to pay for an environmental improvement following a natural disaster.[11] The valuation questions were preceded by an information booklet and a detailed discussion of the valuation context. First, individuals were asked to assume that the Endangered Species Act was not reauthorized and, thus, there would no longer be a Federal

requirement or funding for species protection and restoration. Second, they were informed of the habitat damage caused by Hurricane Hugo. The probability of population survival was given for several population sizes and management practices: (1) with a short rotation plan for the longleaf pine ecosystem at the post-hurricane population size, the population would have a 0% chance of survival; (2) with a long rotation plan which allows for "old growth", at the post-hurricane population level, the population would have a 50% chance of survival; and (3) with a long rotation in combination with specific recovery activities, the population would be expected to grow in size and would have a 99% chance of survival. While introducing uncertainty into the CV scenarios creates complexity for respondents, this was the most accurate way to portray conservation options. As Montgomery et al. (1994) argue: "Because species survival is not a certainty under any conservation program, one can speak only of the likelihood of survival over some time period. Thus the appropriate 'output' unit for both benefit and cost functions is the probability of species survival" (1994, p. 112).

Given the contingent setting, individuals were asked how much they would be willing to pay into an independent foundation whose sole purpose would be the protection and restoration of the red-cockaded woodpecker.[12] Multiple WTP questions were asked to increase the amount of management information provided by the survey. For the woodpecker and habitat restoration activities on the FMNF, the WTP questions followed a natural sequence, which minimized the potential for question order bias (Mitchell and Carson 1989, p. 258). The first question asked for willingness to pay each year to protect the remaining old-growth longleaf pine resource which would maintain the existing probability of RCW survival (50%) and protect the population from certain extirpation. The second question asked for willingness to pay to finance specific restoration activities, which would improve the population's chance of survival from 50% to 99%.[13]

Asking the same individual a sequence of questions has the advantage of increasing the statistical efficiency of the sample information relative to having different individuals answer one question for each of the protection levels. The results in this paper focus on the second question. While the contingent scenario asks respondents to assume that the ESA is not reauthorized, in reality, the ESA does require that restoration activities be undertaken for endangered species. Assessment of the second contingent valuation question provides insight into the value of restoration activities that will actually be undertaken, and thus is of most interest to forest managers.

Following accepted CV protocol (Carson 1991; Mitchell and Carson 1989; Desvousges et al. 1984), the survey instrument underwent several changes as the result of four focus groups and a two hundred person mail pretest in four states.[14] The goals of the open-ended pretest were to test the survey instrument, to determine appropriate dollar value ranges to use in the full survey, and to obtain some preliminary estimates of value for RCW population restoration. After the initial mailing, there was a single postcard follow-up. The pre-test response rate was 21%.

Open-ended willingness to pay values in the pretest ranged from $0 to $100. Since the pretest involved a small number of responses, and since some researchers have had problems with a number of respondents saying yes to their highest bid amount (see Desvousges et al. 1993), the range was widened for the PC and DC versions. PC values included the following dollar amounts: 0, 1, 2, 3, 5, 10, 15, 20, 25, 50, 75, 100, 150, 200, and 300. The dollar amounts were chosen to approximate a lognormal distribution. A considerable amount of research in stated preference methodology has focused on optimal bid design (eg., see Boyle et al. 1998 for a review), and evidence supports having a larger number of choices at lower dollar values and fewer choices at higher dollar values (Rowe et al. 1996). Respondents were also given the opportunity to write in some other dollar amount. DC values were randomly assigned to survey respondents and included the following dollar amounts: 1, 5, 10, 25, 50, 100, and 150. The DC question was a double-bounded (DB) question: if a respondent said yes (no) to the initial bid, he or she was presented with a second bid that was twice as high (half as much) as the original bid. Therefore, the DC bid amounts ranged from $0.50 to $300.

After making modifications to the survey instrument based on results and comments in the pretest, the survey was sent to 1,500 individuals, with 500 receiving the open-ended format, 500 the payment card format, and 500 the dichotomous choice format. The survey was mailed to 750 people randomly chosen in South Carolina, where the primary good being valued was located, and to 750 people randomly chosen throughout the other states of the US.[15] Accepted survey protocols were carefully followed (Dillman 1978), and the final response rate was 51.4%.[16] Of all responses, 33% were from individuals who received the open-ended version, 34% from those receiving the payment card version, and 33% from those receiving the dichotomous choice version of the survey.

4. Results

Table I presents the socioeconomic characteristics of the sample. When compared to statistics for the United States, median household income and years of education are somewhat higher for the sample. Although there are some differences between the sample and population averages, these differences should not affect the results of the comparison tests, since treatments were randomly assigned.

Table II shows the frequency of occurrence of various bid amounts across question formats. In the payment card version, only one individual chose the highest presented offer amount ($300) as his/her willingness to pay. This closely corresponds to the one individual who wrote in a value of $240 in the open-ended survey. As expected, a declining percentage of respondents said yes to the dichotomous choice bid amount as that bid increased. Contrary to a number of recent studies that found a "fat tail" in the distribution of WTP values (Desvousges et al. 1993; McFadden and Leonard 1993; Ready et al. 1996), this pattern did not occur in the

Table I. Socioeconomic characteristics of the random sample and US population.

Characteristic	Random sample	US population
Median income	37,500	29,943
Median schooling	14	12.4
Mean family size	2.7	2.63
% White, no Hispanic	86	80
% African American	7	12

data in this study. No respondents said yes to the highest or second highest DC bid amounts. Offered amounts choked off demand at the highest offered prices.

The first comparison is a test of homogeneity of proportions for the OE and PC question formats using the data presented in Table II. For the test to be applicable, every amount category (cell) should have an expected frequency of at least five. To achieve this expected frequency for each cell, some amount categories are combined and result in the following bid categories: 0–4, 5–9, 10–14, 15–24, 25–29, and greater than 30. The hypothesis of equal proportions is rejected at the five percent level of significance,[17] and it can be concluded that at least one bid category differs for the OE and PC question formats. An analysis of the data indicates that the OE format contains more zero bids, while the PC format has more bids at the five dollar level. Furthermore, the PC format offered respondents the opportunity to circle bid amounts of two or three dollars, while no one wrote in those amounts on the OE survey: OE values tend toward common dollar increments (1, 5, 10, 25, 50). In testing for robustness, hypothesis test results are sensitive to the choice of categories. If the 15–24 and 24–29 categories are combined, the result is significant only at the ten percent level. A third break-down of categories (0–3, 4–9, 10, 11–20, 21–29, greater than 30) yields no significant difference across question formats. Thus, while the distributions of the OE and PC data are similar, they are not identical.[18]

Response rates by CV question format are shown in Table III. The lower response rate for DC is the opposite of what might be expected *a priori*. However, a hypothesis test of equal proportions cannot be rejected, and it is concluded that there is no significant difference in response rates across question formats. This holds true for either of two definitions of response rate.

Respondent difficulty associated with answering valuation questions can be evaluated by examining the rates of item non-response and protest response for the different question formats. As shown in Table IV, the double-bounded dichotomous choice item non-response rate is significantly higher than that for the payment card at the five percent level of significance. This is somewhat contrary to expectations, since the dichotomous choice technique is argued to be the easiest for respondents. The higher non-response rate for the DC format might be explained by the fact that

Table II. Frequency of bid amount by CV question format.

Bid amount	OE frequency[a]	PC frequency[a]	DC frequency[a,b]	
			Yes	No
0	132	117	—	—
1	5	9	12	9
2	—	5	—	—
3	—	4	—	—
5	8	26	4	18
8	1	—	—	—
10	14	17	4	17
15	1	7	—	—
20	4	5	—	—
25	10	6	3	22
30	1	—	—	—
35	1	—	—	—
40	1	—	—	—
50	6	3	1	20
75	0	0	—	—
100	2	2	0	26
150	—	—	0	33
200	1	1	—	—
240	1	—	—	—
300	—	1	—	—
Survey returned				
Unanswered	24	22	42	
Item				
Non-response	13	9	17[c]	
Protest bids	34	19	40	
Total returned	225	234	223	
Total mailed	500	500	500	

[a]OE = open ended, PC = payment card, and DC = dichotomous choice.
[b]Responses to the initial bid.
[c]For double-bounded. There were 12 non-responses to the single-bounded.

Table III. Comparison of response rates by CV question format.

Question format	Response rate 1[a]	Response rate 2[b]
Open ended	51.98	50.89
Payment card	52.08	50.60
Dichotomous choice	50.00	46.53
Test for equal response rates:		
$z_{OE,PC}$	–0.0298[c]	0.0825[c]
$z_{PC,DC}$	0.6067[c]	1.1557[c]
$z_{OE,DC}$	0.5671[c]	1.2202[c]

[a]Response rate 1 = (# Completed + Unable + Refused) / # Delivered.
[b]Response rate 2 = # Completed / (# Delivered – Unable – Refused).
[c]Not significant at the 10% level.

Table IV. Comparison of item non-response and protest bids by CV question format.

Question format	% Non-response	% Protesters
Open ended	6.4677	18.085
Payment card	4.2453	9.360
Dichotomous choice (DB)	9.3923	23.669
Dichotomous choice (SB)	6.6298	same as DB
Test for equal item non-response and equal protest response:		
$z_{OE,PC}$	1.01	2.518**
$z_{DBDC,PC}$	2.05**	3.762***
$z_{SBDC,PC}$	1.05	NA
$z_{DBDC,OE}$	1.06	1.299
$z_{SBDC,OE}$	0.06	NA

**Significant at the 5 percent level.
***Significant at the 1 percent level.

a double-bounded DC question was asked, and respondents may tire when faced with iterative questions. If one compares the non-response of just the first dichotomous choice question, the difference between the single-bounded dichotomous choice and the payment card non-responses is no longer significant. Table IV also shows that the rate of protest response[19] is significantly higher for the DC and OE formats relative to the PC format.[20] Taken together, these results suggest that the payment card format may ease the valuation task faced by survey respondents and lead to efficiencies in data collection.

Table V. Comparison of mean willingness to pay for red cockaded woodpecker habitat restoration by CV question format.

	Protesters deleted			Protesters included		
	mean	s.e.	n	mean	s.e.	n
OE	10.28***	2.35	154	8.42***	1.95	188
PC	8.35***	2.14	184	7.57***	1.95	203
DC	13.25***	2.20	128	10.29***	1.62	168

Test for equal means:

	t-statistics	t-statistics
OE vs. PC	0.608	0.308
OE vs. DC	0.909	0.720
PC vs. DC	1.550	1.040

***Significantly different from zero at the 1 percent level.

Results of three pair-wise t-tests of equal mean WTP are shown in Table V. The reported OE and PC means are univariate means, the simple averages of the values reported by respondents in the survey, and DC mean values are computed using the parameter estimates from a logit regression of response on bid amount, utilizing the likelihood function presented by Hanemann et al. (1991):

$$\text{MeanWTP} = \beta^{-1}\ln(1 + e^{\alpha})$$

Results are reported for models including and deleting protesters. The results presented here hold true whether or not protest bids are deleted: (1) all mean WTP values are significantly different from zero, and (2) there are no significant differences in mean WTP values across question formats. This second result lends support to the convergent validity of the contingent valuation method. Furthermore, the sample medians are equal ($0).

In previous pair-wise comparisons of mean values using different question formats (elicitation techniques), results have shown means to differ significantly, often by orders of magnitude. The question then arises: why is no statistical difference found in this particular study? One potential explanation is that the PC and DC values were well-chosen based on information gained from the focus groups and pre-test. There is close correspondence between the OE and PC responses, as well as consistency across these values and the responses to the DC bid amounts. However, some may argue that the lack of statistical difference may be due to an insufficient number of observations (and thus insufficient statistical power). To investigate this potential lack of power, and to analyze the data with a single estimator, all observations were pooled into a single dataset, and further analyses were conducted.

Table VI. Maximum likelihood parameter estimates using actual and simulated double-bounded dichotomous choice data.[a]

Variable	Protesters deleted		Protesters included	
	Estimate	Std error	Estimate	Std error
Constant	–0.596***	0.177	–0.652***	0.170
Bid Amount	–0.025***	0.005	–0.023***	0.004
OE Dummy	–0.303	0.213	–0.532**	0.214
PC Dummy	–0.305	0.221	–0.433**	0.204

[a]Pooled sample size = 467 with protesters deleted, 560 with protesters included. The number of iterations = 500.
**Significant at the 5 percent level.
***Significant at the 1 percent level.

Table VI shows the results of Monte Carlo simulations for the pooled synthetic dataset. Right hand side predictors include a constant, the bid amount, and a dummy variable for both the OE and PC datasets. Results are sensitive to the treatment of protest responses. For the model that excludes protest bids, parameter estimates on the OE and PC dummy variables are not significantly different from zero, indicating that question format does not influence WTP responses. In contrast, the OE and PC dummies are significantly different from zero (at the 0.05 percent level) when protesters are included in the model estimates. This result implies that: (1) one minus the empirical cumulative distribution functions (*cdf*) for the OE and PC data lie below one minus the *cdf* for the DC data, and (2) the treatment of protest responses can have a significant effect on the convergent validity of WTP estimates.

Mean WTP values for restoration of a red cockaded woodpecker population in South Carolina ranged from a low of $7.57/person/year (payment card, protesters included) to $13.25/person/year (double-bounded dichotomous choice, protesters deleted). Although very few studies have estimated the value of protecting endangered species, these estimates can be compared with values from the studies by Bowker and Stoll (1988) and Stevens et al. (1991) which elicited WTP values for protecting whooping crane and bald eagles from extinction, respectively. Similar to this study, they used payments to a foundation as the payment vehicle, and both studies used dichotomous choice questions for eliciting WTP. Bowker and Stoll (1988) report WTP estimates from $21–$42 for the linear logistic model estimated from their mail sample. Using a log-logistic model (which tends to increase estimates of WTP), Stevens et al. (1991) report WTP values of $28.25 (also for a mail survey). While the WTP estimates from these two studies appear to exceed the values for protecting RCW (roughly by a factor of two to four), it should be recalled that the RCW values reported in this paper are for increasing the probability of population survival from 50% to 99%. Because the other studies did not characterize the extinction probability for the *status quo* program (i.e. do

nothing), their results are not directly comparable to the results of this study.[21] Furthermore, the bald eagle and whooping crane are both more widely known species than the red-cockaded woodpecker, another factor that would be expected to lead to higher WTP values for those species.

5. Concluding Comments

Which, if any, CV question format to use remains an unresolved issue. While the NOAA panel (Federal Register 1993) endorsed the DC format for its ease of use and resemblance to every day decision-making, this study finds that double-bounded DC has a lower response rate (although not statistically significant), a higher item non-response and higher level of protest response than does either the OE or PC format.[22] In contrast, the PC format is found to have desirable response properties.

The convergent validity of CV formats remains an important issue. In this study, convergent validity is found in a three-way comparison of nonuse values for endangered species restoration following a natural disaster. This supports the conclusion reported by Loomis et al. (1997) for a private good (an art print) and by Boyle et al. (1996) for nonuse values to avoid environmental degradation (oil spill containment and cleanup), but is in sharp contrast to conclusions regarding nonuse values reported by McFadden (1994) for Western wilderness areas, Kealy and Turner (1993) for acidification in Adirondack lakes, and Desvousges et al. (1993) for migratory wildfowl. Given the limited research experience with applying CV formats to evaluate nonuse value, it is important to consider why the results of this study indicate validity for this important class of goods.

The nature of the valuation problem (restoring an environmental good that has been damaged by a natural disaster) and the fact that the environmental disturbance was widely known suggest that the necessity for the restoration program was likely viewed as reasonable by survey respondents. The McFadden (1994) and Kealy and Turner (1993) valuation problems were framed in a damage prevention context, which may have affected respondents' perceptions of the need for such programs.

A recent Monte Carlo simulation study by Huang and Smith (1998) suggests that models based on dichotomous choices where WTP is dominated by passive use values (such as this study of an endangered species) are likely to have smaller specification errors than where large use values influence decisions. They argue that specification errors might explain differences found in prior comparative analyses of the effects of question format on WTP.

This research provides evidence that the CV method can elicit WTP values that satisfy convergent validity for goods with significant nonuse value. The conditions under which this conclusion is true deserve closer scrutiny.

Appendix: Willingness to Pay Questions

A. Open-Ended question format:

1. What is the most you would be willing to pay into the recovery fund *each year* to restore the Francis Marion National Forest and to manage the longleaf pine protection area under the long rotation plan? You would be paying to maintain the red-cockaded woodpecker population at level B and to prevent level A (that is, **YOU WOULD BE PAYING TO KEEP THE CHANCE OF SURVIVAL AT 50% AND TO PREVENT IT FROM DROPPING TO 0%**). (please fill in the blank)

I WOULD BE WILLING TO PAY $_____.

2. What is the most you would be willing to pay into the recovery fund *each year*, over and above the amount stated in question 1, to undertake restoration activities for the red-cockaded woodpecker in addition to managing the area under the long rotation management plan? You would be paying to go from level B to level C (that is, **YOU WOULD BE PAYING TO IMPROVE THE CHANCE OF SURVIVAL FROM 50% TO 99%**). (please fill in the blank)

I WOULD BE WILLING TO PAY $_____.

B. Payment Card question format:

1. What is the most you would be willing to pay into the recovery fund *each year* to restore the Francis Marion National Forest and to manage the longleaf pine protection area under the long rotation plan? You would be paying to maintain the red-cockaded woodpecker population at level B and to prevent level A (that is, **YOU WOULD BE PAYING TO KEEP THE CHANCE OF SURVIVAL AT 50% AND TO PREVENT IT FROM DROPPING TO 0%**). (please circle one amount)

$0	$1	$2	$3	$5	$10
$15	$20	$25	$50	$75	$100
$150	$200	$300	Other $_____		

2. What is the most you would be willing to pay into the recovery fund *each year*, over and above the amount stated in question 1, to undertake restoration activities for the red-cockaded woodpecker in addition to managing the area under the long rotation management plan? You would be paying to go from level B to level C (that is, **YOU WOULD BE PAYING TO IMPROVE THE CHANCE OF SURVIVAL FROM 50% TO 99%**). (please circle one amount)

$0	$1	$2	$3	$5	$10
$15	$20	$25	$50	$75	$100
$150	$200	$300	Other $_____		

C. Dichotomous Choice question format:

1. Would you be willing to pay $_____ into the recovery fund *each year* to restore the Francis Marion National Forest and to manage the longleaf pine protection area under the long rotation plan? You would be paying to maintain the red-cockaded woodpecker population at level B and to prevent level A (that is, **YOU WOULD BE PAYING TO KEEP THE CHANCE OF SURVIVAL AT 50% AND TO PREVENT IT FROM DROPPING TO 0%**). (check one)

[ ] YES → If YES, would you be willing to pay $_____? (check one) [ ] YES

[ ] NO

[ ] NO → If NO, would you be willing to pay $_____? (check one) [ ] YES
[ ] NO

2. Would you be willing to pay $_____ into the recovery fund *each year*, over and above the amount stated in question 1, to undertake restoration activities for the red-cockaded woodpecker in addition to managing the area under the long rotation management plan? You would be paying to go from level B to level C (that is, **YOU WOULD BE PAYING TO IMPROVE THE CHANCE OF SURVIVAL FROM 50% TO 99%**). (check one)

[ ] YES → If YES, would you be willing to pay $_____? (check one) [ ] YES
[ ] NO
[ ] NO → If NO, would you be willing to pay $_____? (check one) [ ] YES
[ ] NO

Notes

[1] The dichotomous choice method is also referred to as the closed-ended, take-it-or-leave-it, discrete choice, or referendum method.

[2] Cameron and Quiggin (1994) find that values implied from first and second dichotomous choice responses are highly correlated and may be drawn from the same distribution, but are not identical. A bivariate probit analysis of the data utilized in our study likewise finds a highly significant correlation between first and second DC responses.

[3] This is in contrast to the OE method, which the panel felt invites strategic overstatement.

[4] The identification of starting point bias requires a reference point for comparison of DC responses. This can be a comparison of initial and follow-up DC responses in a DBDC format as shown by Herriges and Shogren (1996), single-bounded DC and OE responses as shown by Boyle et al. (1997), or single-bounded DC and PC responses as shown by Holmes and Kramer (1995). While this paper does not specifically test for starting point bias, this is an important area for future research.

[5] Comparing value estimates derived from different question formats provides a basis for tests of convergent validity (Carmines and Zeller 1979; Mitchell and Carson 1989).

[6] Critics could still argue that, even though all formats are equivalent, no format is accurate.

[7] This would be true if respondents became frustrated with the willingness to pay questions and decided not to return the survey.

[8] To estimate mean WTP values for the double-bounded dichotomous choice data, the likelihood function presented by Hanemann et al. (1991), as well as the code for generating bootstrap estimates of mean and standard deviation, were programmed in LIMDEP (Greene 1992). WTP was computed as $\beta^{-1} \ln(1 + e^{\alpha})$.

[9] The Krinsky-Robb (1986) procedure (KR), which is based on resampling from the estimated variance-covariance matrix and is a fast procedure relative to other methods that resample the raw data (Duffield and Patterson 1991) or regression residuals (Freedman and Peters 1984; Kling and Sexton 1990), is used. The KR bootstrap is also used by Park et al. (1991) to estimate WTP confidence intervals from (single-bounded) DC data.

[10] PC data are actually interval data, but by categorizing the OE data, relevant comparisons can be made.

[11] Other CV studies have attempted to elicit WTP value *in case of* some environmental degradation (see, for example, Boyle et al. 1996). Few studies have measured WTP for improvements following degradation, and fewer still focus on damages that are the result of a natural disaster as opposed to human action.

[12] Different payment vehicles were tested during focus groups. Respondents overwhelmingly preferred the payment into a preservation fund as opposed to paying higher prices for wood and

paper products. There was concern about constructing a realistic tax scenario for funding a relatively low cost public good. Similar to Bowker and Stoll (1988), Brown et al. (1996), and Stevens et al. (1991), the study's payment vehicle is a preservation fund.

[13] The two CV questions are given in the appendix. Because every respondent answered the question about preventing a decrease in the chance of survival before the question about increasing the chance of survival, individuals may have anchored on their response to the first question. However, asking the payment questions in reverse order would have changed the natural path of woodpecker protection activities and confused respondents (i.e. programs that protect woodpecker populations at current levels must occur before programs can be put in place to enhance population survival). Also, payment questions for the enhanced level of survival were framed in terms of WTP over and above the amount in the preceding question.

[14] The mailing lists for both the pretest and the full survey were provided by a professional marketing firm.

[15] This stratification was undertaken in keeping with suggestions in the literature (Kish 1965; Sudman 1972; Kalton and Anderson 1986) to sample more heavily in areas where a rare trait (here, a positive valuation for the good in question) is most likely to occur. However, there were no significant differences in mean values across the two regions. For this reason, and because the primary purpose of this paper is to compare question formats, data from the two regions are aggregated.

[16] This response rate is in the middle of the range (22–79%) of response rates reported in a review of CV mail surveys of general populations (Dalecki et al. 1993). A leading social survey researcher suggests that for mail surveys in general, response rates of at least 50% are adequate for analysis (Babbie 1995).

[17] The value of the chi-squared test statistic is 15.18, which exceeds the critical value of 14.067 (given seven degrees of freedom and alpha equal to 0.05).

[18] Since the results are sensitive to the number of observations in each cell, and are sensitive to the aggregation strategy, this does not provide a definitive test. However, this result is corroborated in the Monte Carlo analysis later in the paper.

19 Protest bids (protest zeroes for the OE and PC formats, and protest "no's" for the DC format) are determined by a follow-up question to the willingness to pay question. Respondents were asked the following for OE and PC formats (for the DC format, $0 is replaced with no, and the answer choices are adjusted to reflect the different question wording): "If you answered $0, please check the reason below that best describes why you answered $0: that is what it would be worth to me; people should not have to pay for the restoration or preservation of the red-cockaded woodpecker or its habitat; I cannot afford to pay for the restoration or preservation of the red-cockaded woodpecker or its habitat; I object to the question; other." Responses two and four were defined as protest bids, since they indicate that the respondent is rejecting the evaluation process (Mitchell and Carson 1989).

[20] The rate of protest response to the DC question in this study is very similar to the protest response rate reported by Stevens et al. (1991) in a dichotomous choice experiment regarding valuation of endangered species.

[21] If one conjectures, however, that the marginal value of RCW restoration increases at a decreasing rate (so that the value of increasing population survival probability from 0% to 99% is more than 2 times the value of increasing population survival from 50% to 99%), RCW values may be no different than either the whooping crane or bald eagle values.

[22] This application is a somewhat novel application of double bounded dichotomous choice in a mail survey, and these results indicate that mail surveys may not be the best use of the DBDC method. In non-mail applications, it may be interesting to determine if the gain in efficiency from adding a follow-up question is offset by the loss in response from having that follow-up.

References

Babbie, E. (1995), *The Practice of Social Research*. Belmont, CA: Wadsworth Publishing Co., 7th edition.

Bowker, J. M. and J. R. Stoll (1988), 'Use of Dichotomous Choice Nonmarket Methods to Value the Whooping Crane Resource', *American Journal of Agricultural Economics* **70**, 372–381.

Boyle, K. J. and R. C. Bishop (1988), 'Welfare Measurements Using Contingent Valuation: A Comparison of Techniques', *American Journal of Agricultural Economics* **70**, 20–28.

Boyle, K. J., F. R. Johnson and D. W. McCollum (1997), 'Anchoring and Adjustment in Single-bounded, Contingent Valuation Questions', *American Journal of Agricultural Economics* **79**, 1495–1500.

Boyle, K. J., F. R. Johnson, D. W. McCollum, W. H. Desvousges, R. W. Dunford and S. P. Hudson (1996), 'Valuing Public Goods: Discrete versus Continuous Contingent-Valuation Responses', *Land Economics* **72**, 381–396

Boyle, K. J., H. F. MacDonald, H. Cheng, and D. W. McCollum (1998), 'Bid Design and Yea-Saying in Single-Bounded, Dichotomous-Choice Questions', *Land Economics* **74**, 49–64.

Brown, T. C., P. A. Champ, R. C. Bishop and D. W. McCollum (1996), 'Which Response Format Reveals the Truth about Donations to a Public Good?' *Land Economics* **72**, 152–166.

Cameron, T. A. and D. Huppert (1991), 'Referendum Contingent Valuation Estimates: Sensitivity to the Assignment of Offered Values', *Journal of the American Statistical Association* **86**, 910–918.

Cameron, T.A. and J. Quiggin (1994), 'Estimation Using Contingent Valuation Data from a 'Dichotomous Choice with Follow-Up' Questionnaire', *Journal of Environmental Economics and Management* **27**, 218–234.

Carmines, E. G. and R. A. Zeller (1979), *Reliability and Validity Assessment*. Beverly Hills, CA: Sage Publications.

Carson, R. T. (1991), 'Constructed Markets', in J. B. Braden and C. D. Kolstad (eds.), *Measuring the Demand for Environmental Quality*. New York: North-Holland.

Cummings, R. G., D. S. Brookshire and W. D. Schulze (eds.) (1986), *Valuing Environmental Goods – An Assessment of the Contingent Valuation Method*. Totowa: Rowman & Allanheld.

Dalecki, M. G., J. C. Whitehead and G. C. Blomquist (1993), 'Sample Non-response Bias and Aggregate Benefits in Contingent Valuation: an Examination of Early, Late and Non-Respondents', *Journal of Environmental Management* **38**, 133–143.

Desvousges, W. H., F. R. Johnson, R. W. Dunford, K. J. Boyle, S. P. Hudson and K. N. Wilson (1993), 'Measuring Natural Resource Damages with Contingent Valuation: Tests of Validity and Reliability', in J. A. Hausman (ed.), *Contingent Valuation: A Critical Assessment*. New York: North-Holland.

Desvousges, W. H., V. K. Smith, D. Brown and D. K. Pate (1984), 'The Role of Focus Groups in Designing a Contingent Valuation Survey to Measure the Benefits of Hazardous Waste Management Regulations.' Draft report to the U.S. Environmental Protection Agency. Research Triangle Park, NC: Research Triangle Institute.

Dillman, D. A. (1978), *Mail and Telephone Surveys: The Total Design Method*. New York: John Wiley and Sons.

Duffield, J. and D. Patterson (1991), 'Inference and Optimal Design for a Welfare Measure in Dichotomous Choice Contingent Valuation', *Land Economics* **67**, 225–239.

Federal Register (1993), 'Natural Resource Damage Assessments Under the Oil Pollution Act of 1990', Department of Commerce, National Oceanic and Atmospheric Administration, 15 CFR Chapter IX. **58**(10), 4601–4614. Friday, January 15.

Fisher, A. C. (1994), 'The Conceptual Underpinnings of the Contingent Valuation Method.' Paper presented at the DOE/EPA Workshop on Using Contingent Valuation to Measure Non-Market Values. Herndon, Virginia, May 19–20.

Freedman, D. A. and S. C. Peters (1984), 'Bootstrapping a Regression Equation: Some Empirical Results', *Journal of the American Statistical Association* **79**, 97–106.

Greene, W. (1992), *LIMDEP Version 6.0: User's Manual and Reference Guide*. Bellport, New York: Econometric Software, Inc.

Hanemann, W. M., J. Loomis and B. Kanninen (1991), 'Statistical Efficiency of Double-Bounded Dichotomous Choice Contingent Valuation', *American Journal of Agricultural Economics* **73**, 1255–1263.

Herriges, J. A. and J. F. Shogren (1996), 'Starting Point Bias in Dichotomous Choice Valuation with Follow-up Questioning', *Journal of Environmental Economics and Management* **30**, 112–131.

Holmes, T. P. and R. A. Kramer (1995), 'An Independent Sample Test of Yea-Saying and Starting Point Bias in Dichotomous-Choice Contingent Valuation', *Journal of Environmental Economics and Management* **29**, 121–132.

Hooper, R. G., J. C. Watson and R. E. F. Escano (1990), 'Hurricane Hugo's initial Effects on Red-cockaded Woodpeckers in the Francis Marion National Forest', *Transactions of the Fifty-Fifth North American Wildlife and Natural Resource Conference*.

Huang, J. C. and V. K. Smith (1998), 'Monte Carlo Benchmarks for Discrete Response Valuation Methods', *Land Economics* **74**, 186–202.

Johnson, R. L., N. S. Bregenzer and B. Shelby (1990), 'Contingent Valuation Question Formats: Dichotomous Choice versus Open-Ended Responses', in R. L. Johnson and G. V. Johnson (eds.), *Economic Valuation of Natural Resources: Issues, Theory, and Applications*. Boulder, CO: Westview Press, Chapter 12.

Kalton, G. and D. W. Anderson (1986), 'Sampling Rare Populations', *Journal of the Royal Statistical Society (A)* **149**, 65–82.

Kanninen, B. J. (1995), 'Bias in Discrete Response Contingent Valuation', *Journal of Environmental Economics and Management* **28**, 114–125.

Kealy, M. and R. W. Turner (1993), 'A Test of the Equality of Closed-Ended and Open-Ended Contingent Valuations', *American Journal of Agricultural Economics* **75**, 311–331.

Kealy, M., M. Montgomery and J. F. Dovidio (1990), 'Reliability and Predictive Validity of Contingent Values: Does the Nature of the Good Matter?' *Journal of Environmental Economic and Management* **19**, 244–263.

Kish, L. (1965), *Survey Sampling*. New York: John Wiley & Sons, Inc.

Kling, C. L. and R. J. Sexton (1990), 'Bootstrapping in Applied Welfare Analysis', *American Journal of Agricultural Economics* **72**, 406–418.

Krinsky, I. and A. L. Robb (1986), 'On Approximating the Statistical Properties of Elasticities', *The Review of Economics and Statistics* **68**, 715–719.

Loomis, J. B. (1990), 'Comparative Reliability of the Dichotomous Choice and Open-Ended Contingent Valuation Techniques', *Journal of Environmental Economics and Management* **18**, 78–85.

Loomis, J. B., T. Brown, B. Lucero and G. Peterson (1997), 'Evaluating the Validity of the Dichotomous Choice Question Format in Contingent Valuation', *Environmental and Resource Economics* **10**, 109–123.

Loomis, J. B., M. Lockwood and T. DeLacy (1993), 'Some Empirical Evidence on Embedding in Contingent Valuation of Forest Protection', *Journal of Environmental Economics and Management* **24**, 45–55.

McFadden, D. L. (1994), 'Contingent Valuation and Social Choice', *American Journal of Agricultural Economics* **76**, 689–708.

McFadden, D. L. and G. K. Leonard (1993), 'Issues in the Contingent Valuation of Environmental Goods: Methodologies for Data Collection and Analysis', in J. A. Hausman (ed.), *Contingent Valuation: A Critical Assessment*. New York: North-Holland.

Mitchell, R. C. and R. T. Carson (1989), *Using Surveys to Value Public Goods: The Contingent Valuation Method*. Washington, DC: Resources for the Future.

Mitchell, R. C. and R. T. Carson (1986), 'Some Comments on the State of the Arts Assessment of the Contingent Valuation Method Draft Report', in R. C. Cummings, D. S. Brookshire and W. D. Schulze (eds.), *Valuing Environmental Goods: An Assessment of the Contingent Valuation Method.* Totowa, NJ: Rowman & Allanheld.

Montgomery, C. A., G. M. Brown, Jr. and D. M. Adams (1994), 'The Marginal Cost of Species Preservation: the Northern Spotted Owl', *Journal of Environmental Economics and Management* **26**, 111–128.

Park, T., J. B. Loomis and M. Creel (1991), 'Confidence Intervals for Evaluating Benefits Estimates from Dichotomous Choice Contingent Valuation', *Land Economics* **67**, 64–73.

Ready, R. C., J. C. Buzby and D. Hu (1996), 'Differences Between Continuous and Discrete Contingent Value Estimates', *Land Economics* **72**, 397–411.

Rowe, R. D., W. D. Schulze and W. W. Breffle (1996), 'A Test for Payment Card Biases', *Journal of Environmental Economics and Management* **31**, 178–185.

Schumann, H. (1994), 'The Sensitivity of CV Outcomes to CV Survey Methods.' Paper presented at the DOE/EPA Workshop on Using Contingent Valuation to Measure Non-Market Values. Herndon, Virginia, May 19–20.

Smith, V. K. and W. H. Desvousges (1986), *Measuring Water Quality Benefits.* Boston: Kluwer.

Stevens, T. H., J. Echeverria, R. J. Glass, T. Hager and T.A. More (1991), 'Measuring the Existence Value of Wildlife: What Do CVM Estimates Really Show?' *Land Economics* **67**, 390–400.

Sudman, S. (1972), 'On Sampling of Very Rare Human Populations', *Journal of the American Statistical Association* **67**, 335–339.

United States Bureau of the Census (1992), *Statistical Abstracts of the United States.*